



Fair Bandwidth allocation between service providers in a geostationary satellite network

Corinne Touati, Eitan Altman, Jérôme Galtier

► To cite this version:

Corinne Touati, Eitan Altman, Jérôme Galtier. Fair Bandwidth allocation between service providers in a geostationary satellite network. RR-4421, INRIA. 2002. inria-00072167

HAL Id: inria-00072167

<https://inria.hal.science/inria-00072167>

Submitted on 23 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***Fair Bandwidth allocation between service providers
in a geostationary satellite network***

Corinne Touati — Eitan Altman — Jérôme Galtier

N° 4421

Mars 2002

THÈME 1



***rapport
de recherche***

Fair Bandwidth allocation between service providers in a geostationary satellite network

Corinne Touati , Eitan Altman , Jérôme Galtier

Thème 1 — Réseaux et systèmes
Projets Mistral et Mascotte

Rapport de recherche n° 4421 — Mars 2002 — 12 pages

Abstract: We consider the problem of how a geostationary satellite should assign bandwidth to several service providers (operators) so as to meet some minimum requirements, on one hand, and to perform the allocation in a fair way, on the other hand. In the paper, we firstly address practical issues (such as integrity constraints), whereafter we provide a computational method for obtaining an optimal fair allocation in polynomial time taking the practical issues into account.

Key-words: bandwidth allocation, fairness, geostationary satellite

Partage équitable de bande passante entre des fournisseurs de services dans un réseau de satellite géostationnaire

Résumé : Considérons le cas d'un système couvert par un satellite géostationnaire dans lequel on souhaite attribuer de la bande passante à différents opérateurs ou fournisseurs de services de façon équitable tout en satisfaisant des contraintes de demandes minimales. Dans cet article, nous présentons tout d'abord les problèmes inhérent à ce type de configuration (comme des contraintes d'intégrité) et proposons un algorithme permettant d'obtenir une allocation optimale en un temps polynomial tout en satisfaisant les contraintes d'ordre pratiques.

Mots-clés : allocation de bande passante, équité, satellites géostationnaire

1 Introduction

We consider a multi-spot geostationary satellite system for which a manager wants to assign bandwidth between various service providers (operators) that operate in different geographical areas. An actual assigned unit of bandwidth may correspond to different amounts of throughput, depending on many factors, and in particular on weather conditions. Indeed, during bad weather in an area, a local operator may have to use a larger part of its bandwidth for redundant information (a higher coding rate for error correction), and thus the effective throughput of information decreases. Therefore, if the objective was to maximize the global throughput, it would become non-profitable to assign bandwidth to operators in areas that suffer from bad weather if this bandwidth could be assigned to other operators instead. It is therefore interesting to understand and then to propose bandwidth assignment schemes that are more fair and do not systematically penalize operators that suffer from bad weather conditions.

The geographical area covered by a geostationary satellite is divided into hexagonal zones called *spots*. Each spot uses a certain frequency range denoted as its *color*. In practice, as the number of available frequencies is limited, each satellite has a fixed number of colors that it can use. Two spots having the same color could interfere with each other. That is why they need to be geographically distant. Therefore, the colors are statically assigned to the spots. The spots are further divided into *zones*. These zones are small enough to assume that the weather condition is the same in any point of a given zone. Then, every operator uses the same coding rate and maximizing the global throughput does not penalize any operator of a zone.

1.1 Technical framework

A central difficulty for solving such systems is that integrity constraints may arise: we might not be able to assign any value of bandwidth between the minimum and maximum given values. Instead, each operator can be assigned one or more carriers in each zone among N types of carriers: $1, \dots, N$. Let B_t be the overall *bandwidth* of a type- t carrier. We assume that $B_1 > \dots > B_N$.

B_t is not directly proportional to the actual throughput of information of a type- t carrier. Firstly, as mentioned before, the throughput depends on the coding rate, which may be different from one zone to another due to atmospheric conditions. Secondly, the effective throughput is lower due to overheads (around 10%) for signaling, frequency margins, etc; the percentage of overhead depends on the carrier type. To handle that, each carrier is associated to the utility $C_t(z, o)$. The utility can be viewed as the value that the operator is willing to pay for a carrier of type t . It can be chosen as a function of the amount of redundancy (in the channel coding) which depends on the atmospheric conditions at each zone; in that way the utility can be made proportional to the actually assigned throughput, so the problem solved becomes how to fairly (or optimally) assign throughput.

We assume that there is a minimum and a maximum number of type- t carriers per zone z required by each operator, denoted by $D_t^{min}(z, o)$ and $D_t^{max}(z, o)$ respectively.

We assume throughout that the minimum agreement can be satisfied. If it not the case, our algorithm can be adapted to find new (feasible) minima, as explained in A.

We allow the manager to assign to an operator more throughput than he or she requires. This means that a type- t carrier can be assigned to an operator that does not need a carrier with greater bandwidth than a j -type carrier for some $j > t$. The operator would be only charged for the price of a j -carrier.

The actual implementation of bandwidth assignment to users involves two phases. The first concerns the allocation of the global bandwidth to each operator in each zone. In that phase we ignore the problem of interferences between carriers of the same colors assigned to neighboring spots. In a second phase, the actual assigned bandwidth has to be translated to physical assigned resources such as slots, if a TDMA approach is used, or such as codes, if the CDMA approach is used. Here we do not address the second allocation problem, which involves complex combinatorial optimization. Yet, in order to simplify that second part of the assignment, we impose an additional constraint on the first phase, namely that **the number of carriers of a type is the same in every spot of a given color**.

1.2 Fairness and global optimization criteria

We are interested here in optimizing a fairness criteria (see [6, 7, 8]). There are several possible ways to understand and implement fairness concepts in our context, in the geographical sense. As we are interested in sharing bandwidth between operators, we may try to achieve:

- local fairness, i.e. in each zone,
- global fairness, i.e. in the total allocation,
- intermediate fairness, i.e. in a geographical area larger than a zone, such as a spot or a group of spots.

We clearly see that two spots having a different color are associated to two independent problems. In the following, we therefore compute the allocations independently among each color group.

Several concepts of fairness can be found in the literature. The most commonly found are the max-min and the proportional fairness. The max-min fairness has been adopted by the ATM-forum [3] as the standard for throughput allocation for the best-effort traffic class ABR (Available Bite Rate). An allocation is max-min fair if and only if there is no possibility to allocate more excess throughput to an operator unless we allocate less excess throughput to another operator that has less excess throughput than the previous one.

It has been shown (in [7, 2]) that the throughput of connections in congestion control protocols for the Internet based on linear increase and multiplicative decrease (such as TCP/IP) achieve a variant of proportional fairness. A bandwidth assignment λ is proportionally fair if for any other assignment λ^* , the aggregate changes in the throughputs is zero or negative. That is to say, if C refers to the set of connections:

$\sum_{c \in C} \frac{\lambda_c^* - \lambda_c}{\lambda_c} \leq 0$. It is shown in [9] that an assignment is proportionally fair if and only if it maximizes the product of the excess of bandwidth.

We may also consider the criterion of maximizing the total used throughput although we saw that it would lead to an unfair solution. It is called *Global Optimization*.

All these criteria produce Pareto optimal solutions ([5]). An allocation λ is said to be Pareto optimal if there is no other allocation that strictly dominates it. This means that there is no allocation where all operators get an allocation greater than or equal to λ with at least one operator getting strictly more than according to λ . The Pareto definition assures that no bandwidth is wasted.

A general fairness criterion that covers all the previous one is due to [9]. Given a positive constant $\alpha \neq 1$, consider the optimization problem:

$$\text{Maximize } \frac{1}{1 - \alpha} \sum_o \left[\sum_{z,t} (D_t(z, o) - D_t^{min}(z, o)) \right]^{1-\alpha}$$

subject to the problem's constraints. Since the utility function is concave and the constraints are linear, this defines a unique allocation which is called the α -throughput allocation. It turns out that this allocation corresponds to the *globally optimal allocation* when $\alpha = 0$, the *proportional fairness* when $\alpha \rightarrow 1$, the *harmonic mean fairness* when $\alpha = 2$, and the *max-min* allocation when $\alpha \rightarrow \infty$.

Similarly, we can define the weighted α -throughput allocation as the unique solution of:

$$\text{Max } \frac{1}{1 - \alpha} \sum_o \left[\sum_{z,t} C_t(z, o) (D_t(z, o) - D_t^{min}(z, o)) \right]^{1-\alpha} \quad (1)$$

We notice by computing the limit of this criterion when $\alpha \rightarrow 1$, that there is no “weighted proportionnal fairness” in the sense of equation 1. This is a remarkable property of proportionnal fairness: multiplying the bandwidth of a connection does not modify the allocation (see [1]).

1.3 Problem formulation

We focus on the fairness criteria obtained as solutions of:

$$\text{Max } \sum_{t,z,o} \frac{[C_t(z, o) (D_t(z, o) - D_t^{min}(z, o))]^{1-\alpha}}{1 - \alpha} \quad (2)$$

This agrees with the previous definition in the case of global optimization. It also agrees with the previous one if we identify a sub-operator to be responsible for all demands of a given type of a given operator at a given zone, and then apply the previous definitions of fairness to the sub-operators. Thus this assignment turns to be Pareto optimal and has essential fairness properties.

Finally, for a given color c , our problem is to compute the optimal J_t , $t = 1, \dots, N$ and $D_t(z, o)$ where J_t is the number of carriers of type t in a spot of color c and $D_t(z, o)$ is the number of type- t carriers assigned on zone z to operator o .

The constraints of our problem can be summarized as follows :

$$\begin{aligned} \forall z, \forall o, \forall t, \quad & D_t^{min}(z, o) \leq D_t(z, o) \leq D_t^{max}(z, o), \\ \forall s \text{ of color } c, \forall t, \quad & \sum_{i \leq t} J_i \geq \sum_{z \text{ zone of } s, o} D_t(z, o),^1 \\ \forall t \in \{1..N\}, \quad & J_t \geq 0, \\ & \sum_{1 \leq i \leq N} J_i B_i \leq B. \end{aligned}$$

2 Our allocation algorithm

We firstly assume that the number of carriers of each type is known and fixed for all spots of a given color, and determine to what operator we should assign them. Then, we show how to determine the number of carriers of each type. Finally, we show how these two components are combined to obtain our optimal fair allocation.

2.1 Assignment of the carriers to operators in a single spot

Within a spot s , let p be the total number of demands, namely, $p = \sum_{z \text{ zone of } s, o, t} D_t^{max}(z, o)$. We consider the set of these demands as a single dimension array (of size p) and note their types t_1, \dots, t_p (where $t_i \in \{1, \dots, N\}$), and values v_1, \dots, v_p (denoted as utility). The demands are sorted so that $v_1 \geq \dots \geq v_p$. Our objective is to maximize the total utility. We will explain in Section 2.3 how to choose the values v_i to obtain the desired fairness criterion.

We work with a fixed given vector $J = (J_1, \dots, J_N) \in \mathbb{N}^N$. J_i corresponds to the number of carriers of type i that can be assigned to demands of type j where $j \geq i$.

We want to find *Accept*, a subset of $\{1, \dots, p\}$, that satisfies the constraint of available carriers (i.e. the number of carriers of type i given by J_i , $i = 1, \dots, N$) and maximizes the value $V = \sum_{i \in \text{Accept}} v_i$.

The algorithm presented in figure 1 solves this problem. The idea is as follows: we add one by one carriers, starting by the most profitable ones. k corresponds to the last level of carrier type that has been filled. More precisely, when all available carriers of type $1, 2, \dots, j$ have been used (the number of assigned carriers of these types has reached $J_1 + \dots + J_j$) then $k = j$. For each new carrier we check whether its type is larger than k . In this case it is accepted and we update the value of k . Otherwise it is rejected.

This algorithm turns out to give the assignment that maximizes the sum V with respect to the vector $J = (J_1, \dots, J_N) \in \mathbb{N}^N$ in a given spot. For a proof of its optimality, see Appendix B.

The algorithm produces a vector "accept" of p boolean elements, and the set *Accept* consists of carriers whose corresponding "accept" value is 1.

2.2 Obtaining a global J

In order to determine the most profitable global $J = (J_1, \dots, J_N) \in \mathbb{N}^N$, we now consider the following global assignment procedure in a given color.

¹This results from the fact that we can assign to an operator a carrier of greater type than it requires.


```

 $k := \min\{i, J_i > 0\} - 1$ 
for  $i:=1$  to  $p$ 
begin
  if  $t_i > k$  then
  begin
    accept $[i]:=1$ 
    for  $l:=k+1$  to  $N$ 
      if  $\sum_{m \leq i \text{ and } t_m \leq l} \text{accept}[m] = \sum_{n \leq l} J_n$ 
        then  $k := l$ ;
    end
  else accept $[i]:=0$ ;
end;

```

Figure 1: An assignment algorithm to maximize the sum of the values.

For each admissible J (satisfying the minimum demand), we find in each spot the best assignment of carriers to operators using the algorithm of Section 2.1. We evaluate the quality of the assignment allowed by J using our global fairness criteria. The J having the best value is selected.

2.3 Basic steps for the general solution

Therefore, the solution of our problem can be found in three steps:

Satisfaction of the minimum requirements. For each spot s we set

$$J_t^s = \sum_{O \text{ operator } Z \text{ zone of } Z} D_t^{\min}(z, o),$$

and set $J^* = \max(J^s, s \text{ spot})$. (The definition and properties of this partial order are given in Appendix 2.1.) If $J^* \in \mathcal{B}$ then all the minimum requirements can be satisfied. Otherwise, the system is non-feasible. From a practical point of view, this step can be done by using the algorithm described in Section C and setting $v_i = 1$ for the minimum demands, and 0 for the other ones.

Maximization of the surplus beneficiaries. Equation 2 shows that, for $\alpha \geq 1$, if for some t, z, o we get $D_t(z, o) = D_t^{\min}(z, o)$, then the optimization criteria given by 2 falls to $-\infty$. Therefore, we aim at minimizing its degree of nullity by maximizing the number of operators to whom we can assign strictly more than their required minimum. For this, we use the algorithm of Section 2.1, and set 2 for the value of a minimum demand, 1 for the value of the smallest demand of each different beneficiary (which can be an operator per zone or an operator per spot or an operator per zone and per type of carrier, etc.), and 0 otherwise. We then restrict our optimization problem only to those operators.

Fair redistribution of the values. We use the algorithm of Section 2.1. We set the value of the minimum demands to ∞ (or any upper bound). If $\alpha \geq 1$, we set the value of the smallest demand to ∞ for the operators to whom we can assign more than their minimum requirements. We then fix the value of each non-assigned demand j ($D_t^{\min} + 2 \leq j \leq D_t^{\max}$) to:

$$\begin{cases} \log \left(\frac{j - D_t^{\min}(s, o)}{j - 1 - D_t^{\min}(s, o)} \right) & \text{if } \alpha = 1^2 \\ C_t(s, o)^{1-\alpha} [(j - D_t^{\min}(s, o))^{1-\alpha} - (j - 1 - D_t^{\min}(s, o))^{1-\alpha}] & \text{otherwise.} \end{cases}$$

If $\alpha < 1$, we use the same equations for all demands above D_t^{\min} ($D_t^{\min} + 1 \leq j \leq D_t^{\max}$)

²One can notice that the value is independent of C . This is a general characteristic of *proportional fairness*: the resource's assignment is independent of the operator's benefit.

3 Numerical Figures

3.1 The considered network

We consider an example with 32 spots and 4 colors, so that 8 spots receive the same color.

Each spot is covered by one MF-TDMA channel (whose bandwidth is 36MHz). It can be used with three types of carriers:

1. Carrier of 2.5 MHz (corresponding to 284 ATM cells/sec).
2. Carrier of 640 kHz (68 ATM cells/sec),
3. Carrier of 161 kHz (17 ATM cells/sec).

If all the bandwidth was assigned to one type of carrier, there could be a maximum number of 14 carriers per channel of the 1st type, 56 of the second type, and 224 of the third type. Under favorable weather conditions, these three carriers correspond to throughput of 2272 kbit/s, 544 kbit/s and 136 kbit/s respectively. This means that if all carriers are of type 2 or 3, then the total throughput is 30 464 kbit/s, and if they are of type 1 then the total throughput is 31 808 kbit/s. The difference is due to a larger overhead due to signaling. An operator may require a 64000 kHz band for a throughput of 136 kbit/s when radio signal fading is more important. Typically the Viterby coding will then be increased from 3/4 to 1/2. Other enhanced signal propagation techniques can also be used.

We suppose that there is 40 operators and that they may use at most 3 spots. Each spot is further divided into three zones. There may be at most two operators per spot, and any operator in a spot may have 1, 2 or 3 zones. The same zone may be assigned to more than one operator

To achieve this, we choose three different spots among 32 for each operator. If one of these spots already has two operators, we simply cancel the request of the new operator. Within each selected spot, the operator chooses at random if it will use each type of carrier. The probability of using a type of carrier is set to 1/2. If it uses no carrier in the end, its request is withdrawn for this spot. We set the maximum number of carriers to infinity. The minimum number of carriers is chosen at random to occupy between 20% and 50% of the bandwidth divided by the amount of types of carriers used by this operator.

For each operator in each spot, we choose a value of C . This value is obtained by multiplying a random (integer) value by the bandwidth of the carrier.

3.2 Numerical results

Results are given by Fig 2. We plot the bandwidth given to each operator in the different spots having the same color, and for all the different types of carriers. Each operator receives a distinct color. In each spot we accumulate the bandwidth assigned to the operators for a particular spot and carrier. The height of the bar is proportional to the bandwidth. Hence, an additional 1-type carrier will increase more the bar than an additional 3-type carrier. We plot the minimum demand plus the surplus, the surplus being on a separate hachured zone.

Fig 2 shows clearly that, as α grows, carriers of higher indexes are favored compared to carriers of lower indexes. Indeed, the fairness criteria redirects progressively the global optimization policy ($\alpha = 0$) to a max-min policy ($\alpha \rightarrow \infty$). Indeed it is more profitable - but less fair - to satisfy the maximum number of demands for a large bandwidth. As α grows, the payload is reduced but small demands receive more attention.

The values of C are summarized in the following figure:

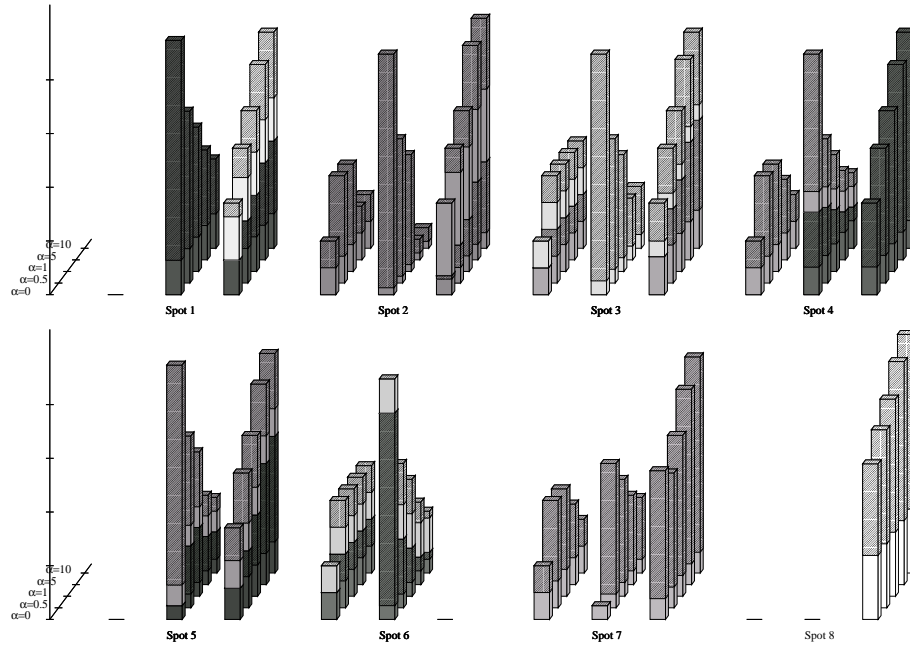


Figure 2: Fair allocation of bandwidth.

S	T	O	C	S	T	O	C	S	T	O	C
1	2	2	1.288	3	0	16	25.00	6	0	4	12.50
1	2	22	1.449	3	1	20	6.400	6	1	18	2.56
1	1	2	2.560	4	2	3	1.449	6	0	18	22.50
2	2	11	0.161	4	0	16	15.00	6	1	4	3.84
2	2	12	0.161	4	1	16	6.400	7	1	17	1.280
2	0	11	12.50	4	1	3	6.400	7	2	17	1.288
2	1	11	5.760	5	2	0	0.161	7	0	17	15.00
3	2	16	0.644	5	2	12	0.322	8	2	28	0.805
3	2	20	0.805	5	1	0	3.200				
3	0	20	22.50	5	1	12	6.400				

S Spot
T Type
O Operator

In spot 7, we can notice that, for $\alpha = 0$ (maximization of the global profit) the available 2-types carriers are used by the 3-types demand instead of the 2-type demand. This is due to a larger value of C (1.288 and 1.280).

In Spot 2, we can observe that, for large values of α although we have 4 available 1-type carrier, only 2 are used for 1-type demand. Indeed, in *max-min fairness* we maximize the revenue of the operator that has the smallest revenue. As $C_3(2, op) = 0.161$ whereas $C_1(2, op) = 12.5$ many more carriers need to be given to the operators that require small bandwidth carriers.

Our algorithm obtains for each value of α a spectrum J that shapes the maximum bandwidth that can be allocated per type of carrier. We plot in figure 3 those values for our series of tests. We can see that the maximum bandwidth allocable to 1-type carrier remains almost constant. A competition for the bandwidth can be observed between the 2-type and 3-type carriers, the smallest carrier having the supremacy in the extreme values of α . The middle values - including proportional fairness and harmonic mean - present favored 2-type carriers. This phenomenon seems to be due to different factors. When α is close to zero, the allocation tries to use the bandwidth as much as possible, independently of the type of carrier chosen. Therefore, small carriers are a good choice since they allow to fill almost all of it with nearly no packing constraint. On the other end, with large values of α , the max-min criteria becomes predominant and constraints the system so much that any additional 2-type carrier becomes useless above a certain limit.

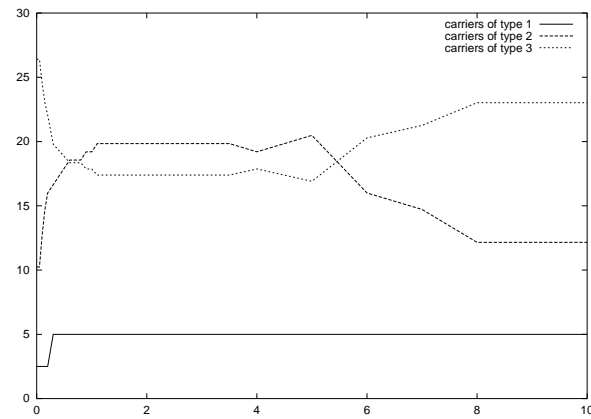


Figure 3: Variation of the bandwidth assigned to the different carriers with α

4 Conclusion

In this paper we have studied different optimality and fairness criteria that can be treated uniformly using a single optimization problem with a parameter α that corresponds to the type of fairness or optimality we wish to achieve. All concepts lead to Pareto optimal solutions. We proposed an efficient computational method to achieve the solution in a polynomial time. The method allows to obtain a large range of solutions depending on the fairness parameter.

References

- [1] C. Touati and E. Altman and J. Galtier, "On fairness in bandwidth allocation", *INRIA Research Report*, September 2001.
- [2] F. Kelly, "Mathematical modelling of the Internet", *Proc. of Fourth International Congress on Industrial and Applied Mathematics*, 1999.
- [3] The ATM forum Technical Committee, *Traffic Management Specification*, Version 4.0, April 1996.
- [4] D. Bertsekas and R. Gallager, *Data Networks*, Prentice-Hall, 1987.
- [5] T. Bonald and L. Massouillé, "Impact of fairness on internet performance", *ACM Sigmetrics*, 2001.
- [6] J. Crowcroft and P. Oechslin, "Differentiated end-to-end Internet services using a weighted proportional fair sharing TCP", *Computer Communications Review*, 28(3):53–67, July 1998.
- [7] F. P. Kelly, A. Maulloo and D. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability", *Journal of the Operational Research Society* **49**, 1998.
- [8] L. Massouillé and J. W. Roberts, "Bandwidth sharing and admission control for elastic traffic", *Telecommunication Systems*, 2000.
- [9] J. Mo and J. Walrand, "Fair End-to-End Window-based Congestion Control". SPIE '98, International Symposium on Voice, Video and Data Communications, 1998.
- [10] H. Yaiche, R. R. Mazumdar and C. Rosenberg, "A game theoretic framework for bandwidth allocation and pricing in broadband networks", *IEEE/ACM Transactions on Networking* **8**, No. 5, pp. 667-677, 2000.

Appendix

A Definitions and computation of fairness for allocation of minimum throughput

In this phase we wish to see whether all minimum required throughput can be allocated.

This can be done by a linear program in which we maximize the total throughputs allocated to all operators, adding the upper constraint that no operator in any zone should get more than its minimum requested throughput. If we denote by λ_s^{min} the minimum throughput guaranteed to operator s and define λ_s the allocated throughput to that operator, then this constraint is given by $\lambda_s \leq \lambda_s^{min}$ for all $s \in S$.

We note that the linear program may end up assigning to some operators their required minimum throughput, whereas others may receive much less than their required minimum. We may consider this as an unfair solution, and prefer a solution in which some more operators get less than their required minimum throughput but the less satisfied operators end up closer to their requirement than in the previously mentioned allocation. In other words, we may try to obtain new values of the minima in a "fair" way. We then may proceed as follows using one of the following two approaches.

1st approach: maximizing throughputs We could use the *same definitions of fairness as before*, set the minimum rate for all operators to zero and the maximum throughput for operator s to λ_s^{min} . Then the computations are the same as before.

2nd approach: minimizing the undelivered bandwidth We could use the *same definitions of fairness as before*, yet instead of applying them to the excess capacity (which we wished to maximize), we can apply them to the amount of bandwidth not delivered $\tilde{\lambda}_s := \lambda_s - \lambda_s^{min}$. We note that this is a non-positive quantity (that we still wish to maximize), and thus we cannot apply anymore some of the formulas for computing the fairness criteria (see, in particular, the one for the proportional fairness).

To avoid that problem we may simply consider the variables $|\tilde{\lambda}_s| = -\tilde{\lambda}_s$ which are positive, and then redefine the fairness concepts accordingly. The max-min fairness criterion will transfer to a min-max criterion in the obvious manner. For the proportional fairness, we can still use the definition with $\tilde{\lambda}_s$ replacing λ_s , or equivalently work with $|\tilde{\lambda}_s|$ and define $|\tilde{\lambda}|$ to be proportionally fair if it is feasible (satisfies the constraints) and if for any other feasible assignment $\tilde{\lambda}^*$, the aggregate of proportional changes is zero or positive

$$\sum_{s \in S} \frac{\tilde{\lambda}_s^* - \tilde{\lambda}_s}{|\tilde{\lambda}_s|} \geq 0. \quad (3)$$

One can then show that the proportional fairness has similar properties as before: it is an assignment that *minimizes* the quantity $\prod_s |\tilde{\lambda}_s|$. Equivalently, it is an assignment that minimizes $\sum_s \log |\tilde{\lambda}_s|$. Unfortunately this problem is computationally more complex than the maximization of these quantities, since unlike the case of maximization, a local minimum may not be a global minimum. Moreover, the uniqueness of the solution is not guaranteed anymore.

A similar complication occurs when redefining the general α -fairness: the maximization has to be changed to minimization and we are not guaranteed to get a unique solution.

We conclude that the first approach is preferable, as it is much easier to solve and it guarantees a unique solution.

B Optimality of the algorithm

The aim of this subsection is to show that our algorithm finds the optimal assignment of carriers for a fixed given vector J .

In the general context, we consider in the following a finite set I of positive valued elements. Let v_i denote the value of element i .

Definition B.1 Let A and B be two disjoint subsets of I . Then define $\text{Union}(A, B, k)$ as a subset H of at most k elements of $A \cup B$ such that $\sum_{i \in H} v_i$ is maximum.

Note that if $I = \{1, \dots, p\}$ and $v_1 \geq \dots \geq v_p$, H can be obtained by the k smallest indices of $A \cup B$.

Definition B.2 Let A and B be two subsets of I . We say that A dominates B iff $|A| \geq |B|$ and there exists an injective mapping φ from B to A such that

$$\forall i \in B \quad v_i \leq v_{\varphi(i)}.$$

Proposition B.1 Let A, B, C and D be three subsets of I such that

- A dominates B ,
- C dominates D ,
- $A \cup B$ and $C \cup D$ are disjoint.

Then, for all $k \in \mathbb{N}$, $\text{Union}(A, C, k)$ dominates $\text{Union}(B, D, k)$.

Proof. Let φ_1 be a dominating maping from B to A , and φ_2 one from D to C . Define ψ from $B \cup D$ to $A \cup C$ as follows:

$$\begin{aligned} \psi(v) &\mapsto \varphi_1(v) & \text{if } v \in B, \\ \psi(v) &\mapsto \varphi_2(v) & \text{if } v \in D. \end{aligned}$$

Obviously the restriction of ψ to $\text{Union}(B, D, k)$ generates a set of at most k elements in $A \cup C$, that is by definition dominated by $\text{Union}(A, C, k)$. \square

Now consider that each element i of I is given a type t_i and denote

$$A_p = \{i \in A : t_i = p\}.$$

Obviously our allocation problem consists of finding a set A such that

$$\forall k \in \{1, \dots, N\} \quad |A_1 \cup \dots \cup A_k| \leq J_1 + \dots + J_k.$$

Theorem B.1 Let $A^{(k)}$ be constructed as follows:

$$\begin{aligned} A^{(0)} &= \emptyset \\ A^{(k+1)} &= \text{Union}(A^{(k)}, I_{k+1}, J_1 + \dots + J_{k+1}) \end{aligned}$$

then $A^{(N)}$ achieves a solution of our allocation problem that dominates all the other solutions.

Proof. Obviously $A^{(N)}$ is a solution of our allocation problem. Let B be an alternative solution. We shall prove by induction that $A^{(k)}$ dominates $B_1 \cup \dots \cup B_k$.

In fact, we have $B_1 \cup \dots \cup B_{k+1} = \text{Union}(B_1 \cup \dots \cup B_k, B_{k+1}, J_1 + \dots + J_{k+1})$. Clearly I_{k+1} dominates B_{k+1} , and by induction $A^{(k)}$ dominates $B_1 \cup \dots \cup B_k$. Hence the result by proposition B.1. \square

C Order among the assignments

We say that an assignment J' is greater than another assignment J (and we write $J' \succeq J$) if and only if

$$\begin{aligned} J'_1 &\geq J_1, \\ J'_2 + J'_1 &\geq J_2 + J_1, \\ &\vdots \\ &\vdots \\ \sum_{t=1}^N J'_t &\geq \sum_{t=1}^N J_t. \end{aligned}$$

Obviously the relation \succeq is a partial order.

We suppose here that an operator has a demand in some spot of a vector of j carriers which has to be satisfied. We assume that by giving more, the operator would be satisfied as well. Hence a demand for j t -carriers is considered to be satisfied when assigning instead j k -carriers, with $k \leq t$.

The advantage of offering more throughput to an operator is illustrated in the following example:

Example C.1 *Spot 1 and 2 have the same color.*

Operator 1 pays 30\$ for a carrier of type 1 in spot 1.

Operator 2 pays 80\$ for a carrier of type 2 in spot 1.

Operator 3 pays 50\$ for a carrier of type 2 in spot 2.

Operator 4 pays 20\$ for a carrier of type 2 in spot 2.

We can satisfy all these demands by a carrier of type 1 and a carrier of type 2 in both spots.

The above "upgrade" scheme implies that a demand J of an operator in a zone, can be satisfied with an assignment J' satisfying $J' \succeq J$.

We have the following key property of the above ordering:

Proposition C.1 *Given two assignments J and J' , there is a unique assignment $J^* = \max(J, J')$ such that $J^* \succeq J$, $J^* \succeq J'$, and for any assignment J'' such that $J'' \succeq J$ and $J'' \succeq J'$, we have $J'' \succeq J^*$.*

Proof We set

$$\begin{aligned} J_1^* &= \max\{J_1, J'_1\} \\ &\vdots \\ J_i^* &= \max\left\{\sum_{k=1}^{k=i} J_k, \sum_{k=1}^{k=i} J'_k\right\} - \sum_{k=1}^{k=i-1} J_k^* \\ &\vdots \\ J_N^* &= \max\left\{\sum_{k=1}^{k=N} J_k, \sum_{k=1}^{k=N} J'_k\right\} - \sum_{k=1}^{k=N-1} J_k^* \end{aligned}$$

which satisfies the required properties. \blacksquare

The interpretation of the above proposition in our context is as follows. Given minimum rate constraints J and J' for two operators, any assignment J'' , that is candidate for being a (feasible) fair assignment, should satisfy $J'' \succeq J^*$. This extends in an obvious way to any number (larger than two) of operators. The value J^* , which can be determined explicitly as done in the proof of Proposition C.1, can now serve as the starting point for fair assignments algorithms: we first determine the unique element J^* and then find the best assignment J'' s.t. $J'' \succeq J^*$.

Unfortunately, the bandwidth assigned to our service is limited. There is therefore an additional constraint on the maximum number of carrier of each type. The constraint is of the form:

$$\sum_{i=1}^N B_i J_i \leq C \tag{4}$$

with C the total bandwidth and B_i the bandwidth of a i -carrier. We denote by \mathcal{B} the set of acceptable assignments with respect to the bandwidth. Obviously if $J \succeq J'$ and $J \in \mathcal{B}$ then $J' \in \mathcal{B}$. Nevertheless, it is possible that for some $J \in \mathcal{B}$ and $J' \in \mathcal{B}$ we have $\max(J, J') \notin \mathcal{B}$. In other words, an assignment can be locally but not globally feasible.



Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
<http://www.inria.fr>
ISSN 0249-6399